心理科学进展 2022, Vol. 30, No. 10, 2321–2337
Advances in Psychological Science
https://doi.org/10.3724/SP.J.1042.2022.02321

机器学习在自闭症儿童早期识别和诊断领域的应用*

侯婷婷[1] 陈 潇[2] 孔德彭[2] 邵秀筠[3] 林丰勋[1] 李开云[1]

([1]济南大学教育与心理科学学院, 济南 250022)
([2]浙江工业大学教育科学与技术学院, 杭州 310023) ([3]青岛市晨星实验学校, 青岛 266000)

摘 要 早发现、早诊断、早干预是开展自闭症儿童教育康复工作的共识, 但传统识别和诊断方法局限及专业人员缺乏常导致自闭症儿童错失最佳干预期。为改善现状, 近年来机器学习凭借其客观准确、简便灵活等方面的优势, 逐渐被应用到自闭症的早期预测、筛查、诊断和评估过程管理中, 积累了较为丰富的成果。但是机器学习也在研究对象选取、分类数据采集和理论模型应用等方面存在局限性。未来研究应推动构建孕产期和新生儿病理生理信息追踪数据库和标准化模型分类指标体系, 同时继续优化算法, 加快智能化自闭症识别和诊断理论成果向实践转化。

关键词 机器学习, 自闭症, 早期识别与诊断, 系统综述
分类号 R395

1 引言

随着智能时代的到来, 人工智能(Artificial Intelligence, AI)与特殊教育的深度融合已经成为教育发展的重要趋势(郭利明 等, 2019)。众多研究者致力于将人工智能技术与不同类型、不同障碍程度特殊教育需要儿童的需求相匹配, 为他们提供更具个性化、精准化的教育康复与医疗诊治服务, 进一步改善其生活和学习质量。随着技术的不断成熟, 自闭症谱系障碍(Autism Spectrum Disorders, ASD)这一数量庞大、病因复杂群体受到了智能教育界的广泛关注, 围绕为自闭症实施行之有效的干预这一焦点问题展开了一系列的探讨, 其中科学准确的早期识别和诊断作为教育干预开展的前提, 不可避免地成为了研究与实践中必须面对和亟待解决的关键问题。

根据美国疾病预防和控制中心数据显示, 2021年自闭症谱系障碍的发病率为 1/44, 与 2000~2002 年的 1/150 相比, 患病率增加了近 240.9% (Maenner et al., 2021)。面对不断攀升的患病率和尚不明确的发病机制, 自闭症已演变成为全球性重大公共卫生问题。积极开展早期预防、识别和干预是应对这一挑战的重要举措。然而目前很少有专门面向 ASD 群体的大范围患病风险筛查, ASD 诊断也主要依靠外在的行为观察与评估人员的主观经验判断。常规诊断流程是评估人员依据标准化诊断工具(如 Diagnostic and Statistical Manual of Mental Disorders - 4/5th ed., DSM-IV/DSM-V; International Classification of Diseases, ICD)对儿童进行评定, 评定方式包括心理和教育测量、医学检查、家长或监护人访谈、日常观察等, 操作费时, 对于评估人员具有极高的专业理论知识和临床经验要求。但现实情况是, 国内具备诊断资质的专业人员非常少, 即使是在特殊教育发展相对靠前的一些国家, 如美国, 能够进行 ASD 诊断的儿科医生也仅占全体医生 8%的比例(陈顺森 等, 2011), 专业诊断“一号难求”。有研究显示, 尽管自闭症儿童的行为与症状早在 1~2 岁就能表现出来, 但典型自闭症诊断的平均年龄却在 4 岁以后, 轻度自闭症的诊断年龄在 5.6~8.6 岁(Sheldrick et al., 2017), 高功能自闭症甚至要延迟到 11 岁以后(Guinchat et al., 2012), 儿童被迫陷入“等待失败”

收稿日期: 2021-11-02
* 教育部人文社科规划基金项目(21YJC880028), 国家自然科学基金项目(32100856)。
侯婷婷和陈潇对本文贡献相当, 两人为共同第一作者。
通信作者: 李开云, E-mail: sep_liky@ujn.edu.cn

的痛苦中。另外这种完全基于外在行为的、主观的诊断纳入的参考变量有限，缺乏客观一致的神经科学指标，误诊、漏诊时有发生(Klin et al., 2000)，一部分儿童被错误地贴上了自闭症的标签，严重违背了评估诊断的初衷。最令人担忧的是，这些问题会严重阻碍自闭症患者早期干预，延误教育康复的黄金时段，给自闭症儿童及其家庭带来巨大的心理和经济负担。最后，即使自闭症儿童诊断完成，后续基于这些患者表现进行的病因、症状研究仍属于回溯性分析，很难全面客观地总结 ASD 行为特征，提取可表征潜在生理机制的标记物，筛选出和临床预测相关度比较高的特征。这使得我们在自闭症儿童防治工作中一直处在被动状态。

机器学习(Machine Learning, ML)作为人工智能的核心，是一类从数据中自动分析获得规律，并利用规律对未知数据进行预测的技术(余明华等, 2017)。该技术自动生成的规则不受个体主观经验的影响，对于提升分类的特异性、敏感性、准确性和增加分类效率都具有重要意义。此外，机器学习可以提取大量数据中细微的、潜在有用的信息，并利用这些信息创建一个针对当前问题的预测或分类模型，特别适用于洞察复杂问题的内在规律，帮助尽早发现可能病因，提高自闭症的生活质量(崔建伟 等, 2021; 李航, 2012; Hosseinzadeh et al., 2021)。国内外学者将其应用于多种精神与神经系统疾病的分类、诊断、治疗和预后管理中，在自闭症群体的预测、识别和辅助诊断实证研究中也取得了大量有价值的成果，但国内在该方面的研究报告相对较少。本文就机器学习在自闭症儿童早期识别和诊断中的应用进行综述，以期为国内相关研究与实践提供参考。

2 研究方法

采用国际上广泛使用的系统性文献综述和元分析方法(Preferred Reporting Items for Systematic reviews and Meta-Analyses, PRISMA)来确定纳入研究的论文构成。该方法包括 4 个文献筛选阶段，共计 27 条综述项目。4 个阶段分别为检索、初筛、纳入和综合，综述项目包括文献标题、摘要、研究目的、研究方法和研究结果等(Moher et al., 2009)。考虑到近年来计算机科学技术发展迅猛，更新换代速度极快，本文选取近 5 年来机器学习在自闭症儿童早期识别和诊断领域文献，具体检索日期范围为 2016 年 1 月 1 日~2021 年 4 月 29 日。应用的英文主题词为："Autis" "Prediction" "Early Screening/ Early Detection/ Early Identification/ Early Recognition" "Early Diagnosis" "Machine Learning/ Artificial Intelligence"，在 EBSCO、Elsevier Science Direct、SAGE、Web of Science、Springer 五个综合性外文数据库进行检索，限定检索范围为同行评议(Peer-reviewed)期刊，获得 520 篇文献，通过参考文献引文检索又获得 50 篇文献，共计 570 篇。作者共同阅读文献的标题、摘要及正文，删除重复、不符合既定标准的论文，最后获取纳入综述的文献 60 篇(见图 1)。

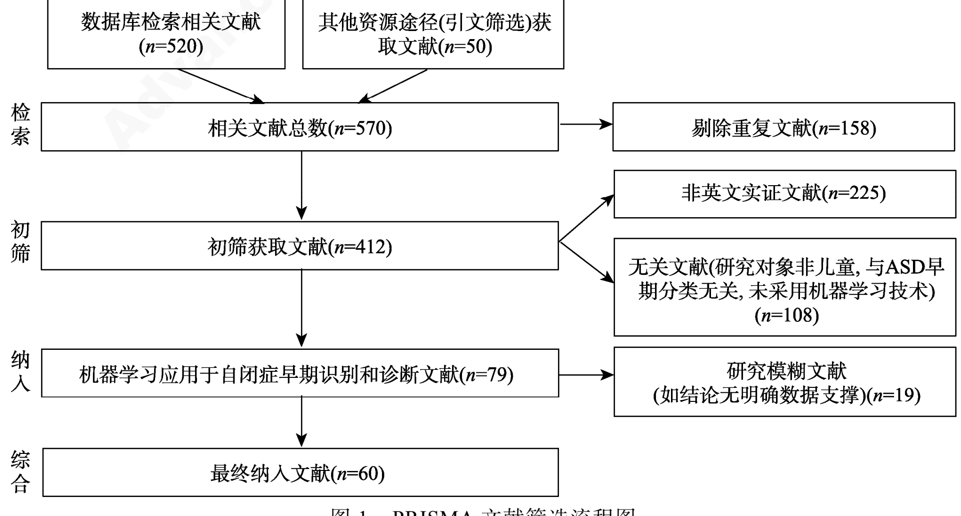


图 1 PRISMA 文献筛选流程图

文献纳入标准：(1)文献为英文实证期刊论文，全文可得且文章不少于 3 页；(2)研究对象中至少包含 1 名 ASD 儿童(在本文中，对儿童的界定参照国际《儿童权利公约》，即儿童指 18 岁以下的个体)，且具有明确的被试统计学变量(年龄)或“儿童”标注[2]；(3)研究包含明确的研究问题、方法和结论，研究结论有详实的数据支撑；(4)研究主题聚焦 ASD 儿童的早期识别或诊断，在摘要或关键词中明确表明采用机器学习技术进行分类。排除标准：(1)非英文文献(韩语、德语、西班牙语等)、综述、书籍章节、会议记录和专利等文章，以及少于 3 页的文章；(2)研究对象全部为 18 岁以上的成年人、老年人或其他特殊群体；(3)研究问题、研究设计或研究结论模糊不清，结论没有数据支撑；(4)与 ASD 无关的精神疾病研究以及未应用机器学习方法的研究。这些标准是在文献搜索之前确定的，以减少在文献筛查期间的潜在偏见。

最后，基于徐建平和张厚粲(2005)提出的“归类一致性公式”检测编码信度，由前两位作者共同计算出纳入综合分析的文献评分者信度在 0.50~1.00 之间，总的归类一致性为 0.812；编码信度系数值在 0.75~0.80 之间，总体编码信度系数为 0.759，表明研究文献具有公平到良好的一致性水平。

3　基于机器学习的自闭症早期识别和诊断研究数据类型与采集

机器学习可以通过对已有数据的抽象，发现数据的隐藏规律(董健宇 等，2020)。因此数据是模型建立的源头性因素，建模的第一步就是数据收集。

3.1　数据构成

ASD 儿童本身异质性高，数据结构性不强，因此分类模型的原始数据一般为二分类数据。从被试构成看，研究者一般同时设置实验组和对照组，选取 ASD 及与其生理年龄、智力年龄等相匹配的其他发展障碍或典型发展儿童(typically developing, TD)，其中障碍儿童常见类型有注意缺陷多动障碍(Attention Deficit Hyperactivity Disorder, ADHD)、发展迟缓(Developmental Delay, DD)、儿童失语症(Childhood Apraxia of Speech, CAS)和其他神经发育障碍(Other neurodevelopmental disorder, OND)。也有研究根据儿童未来患自闭症风险的程度分为高风险(High Risk, HR)和低风险(Low Risk, LR)组。从被试年龄范围看，涉及个体发展的各个阶段，筛查和诊断的最早时间基本上在 3 岁以内，处于婴儿期和幼儿期阶段。自闭症风险预测研究时间更早，一般在 1 岁之前，也有研究提前到孕产期和新生儿时期，但这类研究很少。从被试来源看，涉及美国、英国、法国、中国、伊朗、新西兰、印度、意大利、韩国、西班牙、澳大利亚等数十个国家和地区，常见的数据采集渠道有自闭症研究公开数据库(见表 1)、医院、特殊教育学校或康复中心和少量的在线搜索引擎(如雅虎问答，谷歌搜索)等。不同研究包含的样本数量差异很大，从几十到上千分布不均，其中样本量超过 300 人的研究数据基本上来自于公开数据库。自闭症研究公开数据库是研究者为帮助不同学科、不同专业工作者开展合作研究设置的开放资源库，其储存的大规模数据为使用机器学习开展自闭症识别和诊断带来极大便利。以使用频率最高的 ABIDE (Autism Brain Imaging Data Exchange)数据库为例，该数据库整合全球多个实验室收集 ASD 和 TD 儿童以脑结构和功能影像为主的研究数据，免费向科研人员提供，现已采集完成两期数据，被试数量超过 2200 名，网站注册研究人员涵盖北美洲、欧洲、非洲、亚洲等多个国家和地区。

表 1　自闭症研究公开数据库

序号	中文名称	英文名称(简称)
1	自闭症脑成像数据交换数据库	Autism Brain Imaging Data Exchange (ABIDE)
2	自闭症遗传资源交换数据库	Autism Genetic Resource Exchange (AGRE)
3	波士顿自闭症协会	Boston autism consortium (AC)
4	自闭症治疗网络	Autism Treatment Network (ATN)
5	SSC 数据库	Simons Simplex Collection (SSC)
6	Simons VIP 数据库	Simons Variation in Individuals Project (Simons VIP)
7	加州大学欧文分校知识库	University of California Irvine Repository (UCI)

[2] 作者注：10 项研究中被试人数多、年龄跨度大，同时存在儿童与成年被试，无法做严格区分。所以附表中部分被试年龄大于 18 岁。

机器学习技术的运用突破了传统回溯性研究的局限，自闭症识别和诊断的年龄得以在一定程度上缩短，通过复杂运算提取的隐藏规律还可以被应用到孕产期和新生儿检查中，帮助及早发现高危因素，为早期干预争得时间，但遗憾的是目前这类研究并不多见。除此之外，机器学习技术使大范围、多指标、跨文化数据收集成为可能，从独立小样本数据到跨地区的自闭症公开共享数据库，有效提高了研究的外部效度，也为更深层次自闭症成因挖掘和成功预防提供依据。但应该注意的是，当前的公开数据库多来自于单位或个人研究，数据采集的目的和标准并不一致，样本信息整合缺乏同质性，训练集与测试集之间可能出现偏差，模型的推广问题还需进一步验证。

3.2 数据采集

样本选定后就可以启动数据采集工作。近年来数据采集方式和类型因技术更新而更为多样，例如基于脑电信号和成像技术采集生理信号数据(Nogay & Adeli, 2020; Pagnozzi et al., 2018)、借助量表和观察、访谈、眼动技术等收集行为数据(Bone et al., 2015)。机器学习识别和诊断是基于可以反映自闭倾向的特征及各种神经标志物来进行判断，目前主要有年龄、性别、种族等人口统计学资料，目光对视、社会性微笑、模仿等典型症状表现，儿童个人及家庭病史等临床资料，以及功能性磁共振成像、结构性磁共振成像、脑电图、超声波等生理信号数据(见表 2)，以及不同数据类型组合而成的多模态数据。

3.2.1 基于影像分类

基于脑成像、超声波成像、视网膜成像等生物标记作为机器学习技术的输入特征可以为自闭症分类提供客观证据，因此在研究中应用最为广泛(Abraham et al., 2017; Caly et al., 2021; Lai et al., 2020)。fMRI 的高时间分辨率有助于对大脑特定区域皮层活动的精准定位。sMRI 空间分辨率较高，能够客观地记录不同个体相同脑组织成分间的差异。有学者提出自闭症分类研究中同时采用这两种技术效果最好。如 Rakić 等(2020)采集了 ABIDE-I 数据库中 817 名 7~64 岁被试 sMRI 和静息态 fMRI (resting state fMRI, rs-fMRI)影像，通过 Fisher 算法对特征向量(feature vectors)降维，降维后的特征向量作为分类器的输入信息，最后采用栈式自动编码器和多层感知机算法进行识别，准确率最高可达 85.06%。Sen 等(2018)的研究也发现同时使用两种技术作为特征输入识别 ASD 和 TD 的效果最佳。但是也有研究显示单独采用 fMRI 或 sMRI 的影像也能够实现自闭症个体的准确识别(Eugenia et al., 2020; Katuwal et al., 2016; Plitt et al., 2015)。Guo 等(2017)收集了 ASD 和 TD 的 rs-fMRI 数据，发现经过特征选择的深度神经网络分类准确率可达 86.36%。脑电造价较低、使用方便，可以为自闭症儿童识别提供脑神经细胞电生理活动反应证据。有研究使用 EEG 技术实现了对 3 个月月龄(Dickinson et al., 2020)和 3~14 岁(Abdolzadegan et al., 2020; Grossi et al., 2017)自闭症儿童识别，表现良好。

表 2 基于机器学习的自闭症识别和诊断研究数据类型

一级维度	具体指标
影像	功能性磁共振成像(functional Magnetic Resonance Imaging, fMRI)，结构性磁共振成像(structural Magnetic Resonance Imaging, sMRI)，脑电图(electroencephalogram, EEG)，超声波(ultrasonic, UT)
量表/问卷	自闭症诊断观察量表(Autism Diagnostic Observation Schedule, ADOS)，自闭症诊断访谈量表-修订版(Autism Diagnostic Interview—Revised, ADI-R)，精神疾病诊断与统计手册(*DSM-IV/ DSM-V*)，国际疾病分类(*ICD-10*)，婴幼儿自闭症观察量表(Autism Observational Scale for Infants, AOSI)，修订版婴幼儿自闭症检核表(Modified Checklist for Autism in Toddlers, M-CHAT)，穆林早期学习量表(Mullen Scales of Early Learning, MSEL)，文兰适应行为量表(Vineland Adaptive Behavior Scales, VABS)，社会反应量表(Social Responsiveness Scale, SRS)，自闭症谱系量表(Autism-spectrum Quotient-10, AQ-10)，儿童行为检核表(Child Behavior Checklist, CBCL)，社交沟通问卷(Social Communication Questionnaire, SCQ)，家长问卷[3]
典型行为	眼动行为，社会行为，刻板行为，姿势控制，上肢运动等
个人特征	父母的病史信息，儿童的个人信息(年龄、性别、惯用手、智商、种族、病史、饮食、睡眠等)
其他	DNA, RNA，代谢物，语音

[3] 作者注：在传统的诊断过程中，ADOS、ADI-R、*DSM-V*、*ICD-10* 等主要用于自闭症临床诊断，但是在本研究参考文献中，研究者也会将其用作筛查指标参数，因此正文未做明确划分。

3.2.2　基于量表/问卷分类

传统标准化测量工具汇聚了临床专家丰富理论知识和扎实的实践经验，是自闭症儿童筛查和诊断的“金标准”，可以为机器学习模型构建提供有力证据。基于传统测验工具进行识别和诊断的形式有两种。第一种是将传统标准化测验所得结果借助机器学习技术进行数据分析，探究分类准确性。Abbas 等在 2018 和 2020 年两项研究中均将 ADOS 和 ADI-R 中常用的行为模块作为分类器输入信息，提出了一种基于机器学习的多模块自闭症评估方法。该方法包含 3 个模块，一是家长问卷，二是儿童家庭录像中标记的关键行为，三是临床问卷。通过随机森林、L2 正则化逻辑回归和梯度提升决策树对 18~72 个月的儿童进行分类，发现当同时使用模块一和二，并设置不确定性结果变量(Inconclusive Results Variant, ≤ 25%)时，分类准确率高于 98%，三模块联合时准确性达 92%。即使是三模块共用，测试总时长也仅在 15 分钟左右。也有研究者基于 AQ-10 (儿童/青少年/成人版)和幼儿自闭症量化检核表(Quantitative checklist for autism in Toddlers, Q-CHAT)，设计移动筛查程序 ASD Tests (Thabtah, 2019) 和 Autism AI (Shahamiri & Thabtah, 2020)对自闭症儿童进行预测，ASD Tests 识别准确率最高可达 99.85%，时间仅需 3~5 分钟。

第二种是机器学习算法通过对传统标准化量表或问卷的项目删减，筛选出最具代表性的问题构建分类器(Levy et al., 2017)。如 Duda, Daniels 和 Wall (2016)通过逻辑回归等算法发现，当从 ADI-R 的 93 个条目中筛选出 7 个关键项目或单独选用 ADOS 第 2、3 维度作为分类依据时，对自闭症儿童分类准确率达到了 99%，测试时间缩短到十几分钟。Duda 和 Ma 等(2016)以及 Duda 等(2017)还以 SRS 量表中的 5 种典型行为和 15 个衍生问题构建分类器，分类性能均达到预期(AUC > 93%, AUC = 89%)。值得注意的是，近一半基于量表分类的文献是同时采用两种或两种以上工具，希望通过不同测验的融合，获得更佳分类效果。

3.2.3　基于行为分类

自闭症儿童典型行为表现在发展早期就可以直接观测到，可以为分类模型构建提供外显指标。社交障碍和重复刻板行为是 ASD 诊断的核心标准，可作为行为指标分类的重要依据。Qiu 等(2020)采用静态面孔范式(Still-Face Paradigm, SFP)将 HR-ASD 和 TD 非社会性/社会性微笑、反抗行为、目光交流和积极社交参与的频率和持续时间作为模型分类指标。Negin 等(2021)构建了 HR-ASD 儿童扩展性刻板行为数据库(Expanded Stereotype Behavior Dataset, ESBD)，通过旋转、拍打手臂、手部动作和头部撞击 4 个动作实现与 TD 的区分。眼动仪可以记录人在处理视觉信息时的眼动轨迹特征，眼动行为是基于机器学习进行自闭症分类的最常用行为指标，具体指标包括注视时长和注视频率。Wan 等(2019)采用 SMI RED250 便携式眼动跟踪系统，记录 4~6 岁的 ASD 和 TD 在观看一段 10 秒钟的女性演讲视频时对不同身体部位的注视时长，并采用支持向量机构建分类器，发现儿童对演讲者口部和身体的注视时长差异可以有效区分 ASD 和 TD (Acc- = 85.1%, Sen- = 86.5%, Spe- = 83.8%)。除此之外，部分自闭症群体存在运动障碍，有研究显示姿势控制和上肢运动等指标也可以作为早期分类的依据，其应用准确性可以达到 90%以上(Li et al., 2020; Wedyan et al., 2019)。

3.2.4　基于多模态数据分类

“多模态”是指表征意义的符号系统或有效手段的多样性(顾曰国, 2013)，多模态数据则是指混合多种分类方法、分类指标实现研究目的，在本研究中则指将量表、影像、行为等单一模态的数据采集方式进行整合，同时采用两种或两种以上方法用于自闭症分类模型构建(Geng et al., 2020)。Kang 等(2020)收集 3~6 岁 ASD 和 TD 的 rs-fMRI 数据和对本民族和其他种族人类面孔反应的眼动指标，结合支持向量机的最小冗余-最大相关(minimum redundancy maximum relevance, MRMR)特征选择方法对被试进行分类，发现脑电和眼动数据同时作为分类器数值输入时，分类准确率达到 85.44%，AUC 超过 93%，远高于仅依照单一行为指标分类 66.9%的准确率。Bosl 等(2018)纵向追踪高低风险 ASD 在 3~36 个月间的脑电数据，在 36 个月时采用 ADOS 测量参与者的自闭症症状严重程度，结果发现 3 个月大婴儿的脑电数据可以预测其 3 岁时的 ADOS 得分。韩保祯(2018)也发现多模态信息融合的脑网络数据对于 ASD 和 TD 分类效果最好。

机器学习应用于自闭症早期分类研究数据类

型从单一模态扩展到多模态，分类指标类型越来越多样；采集方法从外显的行为指标到内部神经标记，技术也愈发客观全面。数据采集方式和类型的丰富可以在源头上助力模型构建，为最佳分类效果的获得提供基础。但我们也应该看到，目前数据采集还存在几个问题需要改善，首先影像技术手段比较单一，主要是 fMRI 和 sMRI, ERP、fNIRS 等技术在该领域的应用较少。其次，传统诊断工具本身是经过系统设计的，每个模块有固定目的，不同机器学习方法中对于项目或模块删减标准不一，得到的结果并不统一，工具删减的依据和标准缺乏规范性讨论。传统工具测量时长是为了保证施测过程更为细致准确，机器学习将时间缩短到十几分钟，是否有可能出现信息遗漏导致结果偏差也是一大问题。最后，基于典型行为的分类指标过于广泛，基于 DNA、RNA 等遗传物质的研究却并不多见，不同研究者关注的重点有很大差异，如何在众多指标中选出最具代表性的一种或几种指标仍然值得商榷。这些都是我们在应用机器学习实现自闭症分类之前需要广泛探讨和论证的问题。只有真正解决源头性问题，才有可能将技术的优势真正落实在自闭症群体身上。

4 基于机器学习的自闭症早期识别和诊断研究算法与流程

4.1 机器学习算法

算法是模型建立的技术性要素。机器学习算法的分类非常广泛(Hyde et al., 2019)，基于学习方式不同可分为有监督学习、无监督学习、半监督学习和强化学习，按照模型结构的深度又分为传统机器学习和深度学习。目前在自闭症早期识别中的应用以传统算法为主，包括支持向量机(Support Vector Machine, SVM)、决策树(Decision Tree, DT)、K 最近邻(K-Nearest Neighbours, KNN)、朴素贝叶斯(Naive Bayes, NB)、回归(Regression)、随机森林(Random Forest, RF)和梯度提升决策树(Gradient Boosted Decision Trees, GBDT)等(Ben-Sasson & Yom-Tov, 2016; Meera et al., 2020; Parikh et al., 2019)。一些改进算法也有使用但数量不多，如功能性随机森林(Functional Random Forest, FRF)、改进的蝗虫优化算法(Modified Grasshopper Optimization Algorithm, MGOA)、多层模糊认知图(Multilayer Fuzzy Cognitive Map, MFCM)、ASD-DiagNet 和规则性机器学习(Rules-Machine Learning, RML)等(Eslami et al., 2019; Feczko et al., 2017; Goel et al., 2020; Puerto et al., 2019; Thabtah & Peebles, 2019)。但近年来随着技术的不断发展，人工神经网络(Artificial Neural Network, ANN)、卷积神经网络(Convolutional Neural Network, CNN)、深度神经网络(Deep Neural Networks, DNN)、前馈神经网络(Feedforward Neural Network, fNN)和多层感知机(Multilayer Perceptrons, MLP)等深度学习新方法也逐渐受到了研究者的关注(Achenie et al., 2019; Eni et al., 2020; Raj & Masood, 2020)。

SVM 是传统机器学习代表性算法之一，特别适用于二分类数据问题(Jun et al., 2019; Li et al., 2018; Rrtico et al., 2016)。Liu 等(2016)的研究采用径向基核函数(Radial Basis Function kernel, RBF kernel)将非线性数据投影到高维空间，使数据实现分离，有效区分了 ASD 和 TD，准确率达 88.51%。但有研究显示，SVM 等传统算法往往直接根据特征对样本进行分类，不进行或只进行一次特征变换和选择，过于依赖上游提供的特征，存在对缺失数据敏感和多分类问题处理过拟合问题，在面对自闭症这一成因和表现复杂群体时，效果存在局限(Namdeo & Singh, 2021; 杨剑锋 等, 2019)。深度学习无需借助上游特征，可以将大量原始数据直接输入模型，通过对数据的多次变换检测计算出抽象特征帮助分类(胡越 等, 2019)，模型的分类性能和泛化能力也较传统方法有所改善(Dubreuil et al., 2020; Li et al., 2019; Quaak et al., 2021)。Salari 等(2019)以及 Sewani 和 Kashef (2020)以 rs-fMRI 数据为分类指标，运用 SVM、KNN、RF 和 CNN 分别构建分类器，发现 CNN (70.22%; 84.05%)的分类准确率高于 SVM (69.35%; 69.1%)、KNN (62.11%; 60.1%)和 RF (59.94%; 65.3%)。Dong 等(2021)以静息态脑电图为分类指标构建的 CNN 模型，分类准确率更是达到了 92.63%。Yang (2020)的研究也得到了同样的结论。此外人工神经网络作为深度学习的基本组成部分，也被证明具有较高的分类准确度。Grossi 等(2017)基于 ANN 搭建了一种新型分类系统，将多比例尺分级组织图与时间压缩的隐函数(Multi-Scale Ranked Organizing Map, MS- ROM/Implicit Function as Squashing Time, I-FAST)算法相结合，

以脑电图像为判别指标, 发现该算法对 ASD 预测准确率高达 100%。但也有研究显示, 传统算法可以通过增加分类器数量进行自我优化。随机森林将 bootstrap 再抽样和决策树算法相结合, 构建树型分类器集合(曹正凤, 2014), 分类速度、准确性和抗噪性能明显高于单一分类器算法, 还可以在一定程度上减轻过拟合问题。Wingfield 等(2020)开发的 ASD 筛查程序, 发现 RF 相较单一的 NB、DT 分类效果明显优化, AUC 达到 98%。

机器学习算法无需明确编码, 可以自动检测数据中存在的模式, 处理复杂的非线性数据, 相比传统识别和诊断方法具有强大的功能优势。近年来在 ASD 分类研究中的应用还呈现出从传统算法向深度学习过渡的趋势, 诊断效率大幅提升。但目前超过 75%的研究使用的仍是传统机器学习算法, 对于深度学习的关注有待提高。此外当前研究中算法众多, 不同研究者使用的算法存在很大差异, 同一算法对不同对象施测结果也有不同, 对于最佳算法的争议一直没有停止。因此未来研究应继续深入探讨如何找出最佳算法, 更大程度发挥技术的价值。

4.2　机器学习效果评价

效果评价是模型构建的指挥棒和导航塔。数据收集和算法选择后, 需要对模型有效性加以验证。机器学习效果评价一般是以专业医师诊断为校标, 计算相关系数或成功预测比例得到的。根据原始二分类数据特征, 一般将分类效果分为正类(Positive)和负类(Negative), 其中真阳性和真阴性是预测良好中的正类和负类, 假阴性和假阳性是预测不良中的正类和负类(见表 3)。具体的模型性能评价指标包括：(1)准确率(Accuracy, Acc-), 是指分类正确的样本数占总样本数的比例。通常来说, 准确率越高, 分类效果越好。计算公式为 Acc- = (TP + TN) / (TP + FN + FP + TN); (2)敏感度(Sensitivity, Sen-), 指正类中自闭症儿童分类正确的比例, 又称召回率(recall)、真阳性率(True Positive Rate, TPR)。计算公式为 Sen- = TP / (TP + FN); (3)特异度(Specificity, Spe-), 指负类中自闭症儿童分类正确的比例, 又称真阴性率(True Negative Rate, TNR)。计算公式为 Spe- = TN / (FP + TN); (4)接收者操作特性曲线(operating characteristic curve, ROC 曲线), 是一种以假阳性率为 X 轴, 以真阳性率为 Y 轴组成的图形, 主要通过计算 ROC 曲线下面的面积(Area Under ROC Curve, AUC/AUROC)对分类性能进行比较。一般来说, AUC/AUROC 越大, 分类效果越好。除上述 4 种常用指标外, 部分研究还采用了阳性预测值(Positive Predictive value, PPV)、阴性预测值(Negative Predictive value, NPV)、错误率(false discovery rate /error rate, FDR)等指标来考察分类器效果。

表 3　ASD 二分类数据的混淆矩阵

		预测	
		ASD	Non-ASD
真实	ASD	真阳性 (True Positive, TP)	假阴性 (False Negative, FN)
	Non-ASD	假阳性 (False Positive, FP)	真阴性 (True Negative, TN)

4.3　机器学习流程

逻辑清晰、完备有序的数据处理流程是模型有效落实的显性表现。机器学习应用于 ASD 早期识别和诊断大致经过以下 4 个步骤: 首先是收集数据阶段, 也就是选择 ASD 分类所需影像、量表、行为等数据指标。其次是原始数据处理阶段。该阶段主要目的是实现特征选择, 通过过采样、欠采样等方法对重复记录、缺失值和不平衡数据进行降噪, 降低计算复杂性和提高分类精度, 形成分类所需数据集(Rahman et al., 2020)。然后就进入自闭症和对照组分类阶段。该阶段通常会采用 k 折交叉验证将整个数据集随机分为 k 份, 其中 1 份作为测试集(testing set), k−1 份用于训练, 其中训练的数据又可以随机分为训练集(training set)和验证集(validation set), 这样交叉验证 k 次, 每一份都可以作为测试集进行测试, 平均 k 次的结果, 得到最终结果。具体流程又可简述为：第一, 输入训练集, 采用机器学习算法进行训练, 构建 ASD 分类模型; 第二, 将验证集输入模型中, 比较模型预测结果和“自闭症”标记数据的区别, 验证模型分类准确性; 第三, 测试集测评模型分类性能。最后获取最佳分类模型。Epalle 等(2021)构建的 DNN 模型就是遵从了这一程序, 借助 ABIDE-I 中的脑成像数据区分 ASD 和 TD。数据处理阶段利用过采样进行特征选择, 使 CC200 集增加 19900 个特征, AAL90 集增加 4005 个特征,

DOS160 集增加 12720 个特征; 在分类阶段, 采用10 折交叉验证(10-fold cross- validation)、5 折分层交叉验证(5-folds stratified cross-validation)和留一法(Leave-one-site-out)构建分类器和评估分类结果, 在每一部分数据中, 设置 70%用于训练, 20%用于验证, 10%用于测试。最后模型验证结果显示该分类器对两组被试的分类准确率可达 78.07%。

机器学习应用于 ASD 儿童分类自动分配训练集和测试集, 经过过滤和反复检测选择和提取数据特征, 发现数据隐藏规律, 从来源上减少了数据偏差, 增加了研究的内部效度。模型有效性的验证范围涵盖 Acc-、Sen-、Spe-、AUC、PPV 和 NPV 等多个维度, 越来越注重对于模型的客观测量, 也在一定程度上增加了研究的外部验证。但是通过文献汇总也能够看出, 目前研究中采用的分类指标组合大相径庭, 指标报告的数量参差不齐, 缺乏较为一致、普遍被大家认可的分类效果指标体系。另外, Falkmer 等(2013)认为自闭症分类模型的准确率应至少达到或高于 80%才可以被视为达到了要求, 但当前部分研究的准确率甚至不足 60%, 距离分类准确率的要求还相差不少(Bussu et al., 2018)。后续研究如何从技术层面上继续优化和规范这些问题, 还需要更深入的探讨。

5 基于机器学习的自闭症早期识别和诊断研究应用

5.1 预测自闭症患病风险因素

自闭症作为一种病因不明的神经发育障碍, 仅仅依靠常规的产前检查无法达到预防目的。机器学习技术可以通过广泛收集影像和生物指标信息, 充分挖掘和汇总信息, 计算得出具备预测功能的数据模型, 实现对自闭症患病风险的有效预测(刘笑晗 等, 2022), 为自闭症病因探索与综合诊疗提供新的方向。

孕产期环境和母体自身健康因素都有可能影响胚胎发育质量, 产前对于孕妇(尤其是高风险孕妇)和胎儿的医学检查与监护可能有助于发现自闭症风险因素。机器学习可以通过对超声波检查、羊膜腔穿刺、X 射线检查、绒毛细胞检查、胎儿镜检查等收集的影像信息和生物标记物信息, 推测 ASD 患病影响因素。如 Caly 等(2021)通过收集胎儿的超声波测量值(包括股骨长度、头围、腹围等)和生物学特征(如 IgG 巨细胞病毒水平等)预测新生儿患自闭症的风险, 采用最小绝对收缩和选择算法(Least Absolute Shrinkage and Selection Operator, LASSO), DT 和极端梯度提升(eXtreme Gradient Boosting, XGBoost)算法构建分类模型, 研究结果显示模型阳性预测值可达 77%。母体免疫性疾病家族史、母体对巨细胞病毒的免疫力、孕晚期胎儿股骨长度、孕晚期胎儿白细胞数量、新生儿心率和性别等 10 项指标可以作为 ASD 预测标记物。基因是决定生命健康的内在因素。基因检测可以提取细胞中的DNA分子信息, 确定新生儿基因类型和缺陷。机器学习技术通过梳理复杂的遗传信息, 可以帮助发现与自闭症患病有关的遗传因素。Bahado-Singh 等(2019)利用 450k 甲基化芯片(Infinium HumanMethylation450 BeadChip)检测 ASD 和 TD 新生儿血斑中的白细胞 DNA, 发现在胞嘧啶基因中 CpG 位点甲基化异常显著, 预测结果 AUC 为 100%。同时研究者还发现自闭症在 EIF4E、FYN、SHANK1、VIM、LMX1B、GABRB1、SDHAP3 和 PACS2 基因中存在表观遗传调控障碍。Oh 等(2017)以及 Sekaran 和 Sudha (2020)研究也证明外周血基因表达谱(gene expression profiles)和 FAM104B、CCNDBP1、H1F0、ZER1 等生物标记能够实现对 ASD 个体的分类和预测, 其准确率分别可达 93.8%和 97.62%。还有学者报告长链非编码 RNA (long non-coding RNA, lncRNA)也能够用于预测新生儿罹患 ASD 风险(AUC = 83.9%; Gök, 2018)。Qiu 等(2020)基于 SVM 构建的预测模型显示, 一碳单位代谢(folate-dependent one carbon metabolism, FOCM)、转硫作用(transsulfuration, TS)中的多种代谢物以及母体孕前和孕期摄入的叶酸含量也可以作为新生儿自闭症风险预测的指标, 该模型预测的准确率达到 96%以上。借助机器学习对 HR-ASD 家族婴儿的大脑功能连接磁共振成像(functional connectivity Magnetic Resonance Imaging, fcMRI)进行追踪研究, 将 ASD 风险预测年龄提前到了胎儿出生后 6 个月(Emerson et al., 2017)。

5.2 筛查高危自闭症

早期自闭症筛查可以在大范围人群中发现高危儿童, 对于改善自闭症儿童社会适应、预防继发性发育障碍和减轻家庭与社会负担具有重要意义(陈光华 等, 2022)。如果筛查工作可以直接由家长或其他熟悉儿童的人员操作, 就可以更早发

现潜在 ASD 患者, 避免错失干预时机。Thabtah 等(2018)借助机器学习模型简化传统量表(CHAT、AQ), 开发了一款适合 ASD 筛查的移动应用程序——ASD Tests, 该工具可以在手机端免费下载使用, 适用范围涵盖婴幼儿期(1~3 岁)、儿童期(4~11 岁)、青少年期(12~16 岁)和成人期(17 岁之后)四个阶段, 每个阶段仅有 10 道测试题, 测试时间在 3~5 分钟以内, 非专业人员也能轻松操作。研究发现该程序对于儿童、青少年和成人阶段 HR-ASD 的筛查灵敏度、特异性和准确率均高于 95%, 分类性能良好(Thabtah et al., 2019)。Raj 和 Masood (2020)基于个体性别、种族、是否先天黄疸等 21 项特征构建自闭症筛查模型, 发现 CNN 在儿童、青少年和成人群体中筛查的准确率分别可以达到 98.30%、96.88%和 99.53%, 能够快速地在大范围人群中找出自闭症患者。除此之外, 神经影像数据和行为表现作为风险识别重要的内外部参照, 也被广泛应用于筛查或诊断模型的构建, 取得了令人满意的效果(Dickinson et al., 2020; Lai et al., 2020; Qiu et al., 2020; Rakić et al., 2020)。

5.3 辅助诊断自闭症

诊断是对筛查出的高危儿童做更全面细致的专业检测。传统诊断方法通过观察儿童外在行为和家长报告方式获取数据, 易受评估人员主观性和报告者回忆偏差影响。诊断过程中存在的等待周期长、误诊等问题也阻碍了后续临床干预的实施。机器学习辅助专业人员诊断可以纳入大量客观分类指标, 在一定程度上减少主观施测弊端, 还能够减少耗时, 降低等待成本。社交机器人是 ASD 智能诊断中的常用工具, 其设计有趣亲切, 通过丰富的面部和肢体语言与受测者沟通, 可以减少自闭症社交恐惧(黄碧玉, 2018)。Ramírez-Duque 等(2019)采用 RGBD 传感器采集 HR-ASD 和 TD 幼儿与 ONO 社交机器人互动的行为特征, 并利用 CNN 构建两类儿童面孔识别模型, 辅助专业人员对 ASD 儿童进行诊断, 发现该模型可以实现 ASD 的自动化识别, 花费时间不超过 40 分钟, 准确性与临床专家诊断结果基本一致。Diehl 等(2012)也发现社交机器人可以作为自闭症诊断和干预的有效工具。影像和传统量表也为机器学习辅助诊断提供了有效证据(Bone et al., 2016)。Tariq 等(2018)以 ADOS 和 ADI-R 为指标构建的模型诊断准确率高于 94%, 灵敏度也超过 90%。Abdolzadegan 等(2020)对 3~12 岁的 ASD 和 TD 脑电图进行分析, 基于 SVM 和 KNN 建立诊断模型, 准确率达到 94.68%。

5.4 管理评估和干预过程

自闭症早期预测、识别和诊断的最终目的是为了给儿童提供行之有效的早期干预。收集自闭症群体评估和干预过程信息, 追踪干预施行效果, 可以为确定儿童生理病理机制和治疗靶点提供思路。然而传统的筛查、诊断、干预和预后再评是几个相对独立和分散的环节, 难以对信息进行比较系统的收集和整合, 也就更难发现和把握数据的内在趋势和隐藏特征, 造成了数据的浪费。基于机器学习的分类一大优势还在于对数据的有效利用, 在原有评估和干预数据基础上构建一体化智能评估和干预管理系统, 增强对自闭症儿童诊断和干预的跟踪式监控(Koumpouros & Kafazis, 2019)。Cognoa (Cognoa ASD Screener/ Cognoa ASD Diagnosis Aid)是一款以机器学习技术为支撑的新型 ASD 评估程序, 兼具筛查、诊断和风险预测功能。该程序主要服务对象为婴幼儿至学龄前期自闭症儿童(18 个月~5 岁), 需要家长完成一份包含 15 道题目的问卷, 并上传两个以上时长 1~2 分钟的儿童日常生活视频(如用餐、游戏时间), 资料上传完成后, 后台自动评分输出风险值, 整个过程约为 15 分钟(Kanne et al., 2018), 每次测评结束程序会自动保存结果。2021 年 6 月, 美国食品和药物监管局(Food and Drug Administration, FNA)对该程序应用于自闭症患者诊断的可行性进行核定, 发现来自 14 家医疗中心的 425 名儿童程序诊断结果与专家结论基本完全一致, 获批准上市使用。基于机器学习技术建立的自闭症图片评估智能管理系统, 还可以记录 ASD 从早期筛查到干预治疗过程中儿童能力变化数据, 获得儿童预后再评信息(Perera et al., 2017)。人工神经网络支持的专家系统, 甚至可以在无预先编程设定的情况下机智应对和解决不确定性问题, 为自闭症诊断和干预提供实时建议, 是临床评估中的“得力助手” (靳宇倡 等, 2022; Negin et al., 2021; Rahman et al., 2020)。

综上所述, 机器学习在自闭症分类中的应用涵盖病因分析, 患病风险预测, 潜在危机人群筛查, 辅助诊断和智能病程管理。这些工作的开展有助于增加从业人员对自闭症病因病程信息的掌

握，实现筛查、诊断、干预和预后再评的动态化管理，节约长期随访所需的人力物力资源，更快找出众多个性因素中的“共性”，提高诊疗的有效性，这也是我们未来临床诊疗可以借鉴的一个思路。

6 机器学习应用于自闭症儿童识别和诊断的优势与不足

6.1 优势

目前机器学习技术已经应用于自闭症儿童早期预测、筛查和诊断，其客观性、便捷性和有效性经过了较为广泛的验证，分类性能在不断的优化和完善中也表现越来越好，部分智能工具也开始推广应用。相较于经验诊断和传统测量，机器学习的融入体现了人工智能技术的一些优势，集中表现在 4 个方面。第一，机器学习获取和整理数据指标的能力更强，有助于构建多模态分类模型。传统分类方法主要是通过对外显行为指标的观测得出结论，参考的变量有限，机器学习可以集合多种类型、不同观测方法获得的数据指标，兼顾内在和外部影响因素，减少主观误差，提高预测和识别结果的全面性和客观性。此外，机器学习模型还能够跟踪评估和诊断过程管理，增加了数据的利用率。第二，机器学习降低了预测和识别的时间成本，提高了测评效率，减轻诊断人员压力。分类模型的建立一方面可以通过大范围筛查，快速发现高风险自闭症患者，尽早引起家长和其他相关人员的警觉；另一方面还可以通过减少测试项目、缩短评测流程等有效压缩诊断时间，降低施测的等待时间，减少专业医师压力。第三，机器学习增加了研究的内外部效度。机器学习的建模过程和计算原理区别于传统方法，可以自动将原始数据分为训练集和验证集，无需数据生成模型和参数估计直接输入性能指标，模型还可以随样本数量累积不断优化，提升了研究的内部效度和外部推广力，也有利于加速模型应用于实践的速度。第四，机器学习可自动提取关键特征，帮助建立风险预警机制。机器学习可以捕捉肉眼无法观测到的细微表情或身势动作变化，减少关键指标的遗漏，还适用于非线性数据，可以从纷繁复杂的数据中心抽丝剥茧，更适合分析 ASD 这一类成因不明、结构复杂的疾病，降低症状恶化的可能。尤其是对于婴幼儿和学龄前期儿童，早发现、早干预能够发挥最大效益，帮助这些孩子及其家庭获得更高的生活质量，减少因障碍引发的各种社会适应问题。

6.2 不足与展望

虽然机器学习在 ASD 儿童识别和诊断中具有一定的方法学优势，相关研究成果也逐步增多，但距离大范围应用还有一定距离，其面临的具体问题与可能的解决方法可总结为以下三个方面。

首先，缺少孕产妇和新生儿自闭症病源追踪研究。遗传学研究结果表明，常见的遗传变化，如基因突变、易位、倒位和拷贝数变异等都可能与 ASD 的产生存在一定程度的相关，患有 ASD 的新生儿拷贝数变异频率明显高于 TD (Sanders et al., 2011)。也有研究报告 ASD 可能是由影响细胞增殖、迁移和许多其他基本过程的致病事件序列导致(Caly et al., 2021)。对自闭症儿童大脑的尸检报告也显示其前额叶皮层神经元异常过剩，还存在脑过度生长和巨脑问题(Tyzio et al., 2014)。Rahman 等(2020)还发现父母的年龄、家庭病史也会增加新生儿患自闭症的风险。这些证据表明自闭症受遗传因素的影响，且这些病症在个体胚胎形成和发展时期可能就已经出现(Caly et al., 2021)。除遗传因素外，孕产妇暴露于危险环境也是自闭症患病的易感因素之一。Choi 等(2016)研究发现，ASD 发病率增加与母体病毒或微生物感染、发热以及免疫系统疾病有关。在孕期服用药物，特别是丙戊酸钠(sodium valproate)，和维生素缺乏可能导致胎儿罹患自闭症。剖腹产、早产和新生儿并发症等问题也会增加 ASD 的发病比例(Cloarec et al., 2019)。这提示我们个体生命早期的内外部环境是导致自闭症出现的关键因素。

理清自闭症个体发展初期病理、病因和病程因素可能会为患病风险预测和防控提供新的契机，但遗憾的是当前文献研究中还很少有借助孕产期数据进行自闭症识别的研究。当前文献中仅有的两篇，Caly 等(2021)采集了孕产妇从孕早期到胎儿出生后 1 天内常规生物标记物和超声测量值来预测自闭症，并与儿童 4~5 岁时诊断结果进行比较，预测准确率达到了 77%。Bahado-Singh 等(2019)采集 249 种新生儿(出生 24~79 小时)白细胞表观基因组标志物，成功预测了后期被诊断为自闭症的儿童(Sen- = 97.5%, Spe- = 100%)。这启示我们如果能在胚胎形成和新生儿发展初期，广泛收集孕产妇及其家庭成员健康信息，关注遗传和环境

中可能的致畸因素, 尤其是较为稳定的生物标记物, 借助机器学习模型庞大的数据梳理功能提取其中的共性因子, 是否就有可能逐步摆脱“被动等候”的局面, 变为“主动防御”。因此未来研究可在遵循伦理原则的基础上, 尝试借助孕期筛查和新生儿体检等环节建立儿童生理病理信息数据库, 收集儿童(尤其是高危家族儿童)从胚胎期到出生期间的过程性信息, 也可借鉴美国儿科学会的做法, 在确保参与者知情同意的前提下于固定时间段(如出生时、出生后 8 个月、18 个月等)进行长期追踪研究(Wan et al., 2019), 获取自闭症个体发展初期病源的原始性资料, 归纳可能的风险因素, 及早规避和预防这一障碍的出现。

其次, 缺少标准化模型分类数据体系。机器学习模型为海量信息处理提供了可能性, 但技术本身对于数据的包容也可能成为一把双刃剑。现有机器学习分类指标虽然形式多样, 但内容繁杂, 缺少清晰明确的分类依据和系统化、规则化的应用效果分析, 测评结果的可验证程度不高。以基于行为的单一模态分类指标为例, 研究者所使用的具体指标包含了重复刻板行为、反抗行为、自我伤害行为、精细/粗大动作、社会交往行为、社会性微笑、共同注意、模仿、眼神变化、表达性和接受性语言等数十种行为表现。不同行为表现又可以细化为更微观的分类指标, 仅眼神变化一项就可以分化出眼神检索(eye scanning)、眼神追随(eye tracking)和眼神接触(eye touching/contacting)等多种范式。但是即使是同一行为表现, 得出的结论也不完全一致(Shic et al., 2014; Wan et al., 2019)。基于量表的分类也出现了这一问题, 研究文献中包含多种分类工具, 不同工具蕴含的信息和施测对象也不一致。即使是同一测验工具, 项目的筛选标准与结论也不完全统一(Emerson et al., 2017)。基于多模态的分类模型, 在融合多元信息过程中混杂了更多与测评关联程度不高的信息, 对测试准确性有一定干扰。多种模态的组合, 同一模态不同分类指标的组合, 为研究者提供了更多分类标准选择的同时, 也带来了更多的测量误差, 不利于研究结果的推广和泛化。

因此模型构建过程中应重视分类指标内部秩序性, 梳理出分类指标选取的标准与依据, 逐步确定出一部分稳定性强、适应范围广、准确度高的分类指标, 构建更具科学性和稳定性的分类指标体系。有研究者提出, 年龄可以作为分类指标划分的一个依据。Liu 等(2016)的研究显示婴幼儿和儿童期 ASD 识别更适用于眼神交流、手势和面部表情等非语言行为特征, 青少年 ASD 则更应侧重于和同伴的语言交流与情感互动。随着年龄增长, 社会性沟通与情感联结成为更重要的分类指标。此外, 大脑的变化经常出现在临床表现之前且具有较高的稳定性(Elsabbagh et al., 2012; Wolff et al., 2012; Wolff et al., 2015), 因此分类过程中应特别重视客观影像指标的应用(Emerson et al., 2017; Hazlett et al., 2017; Itani & Thanou, 2021)。言语智商、自闭症严重程度等经过验证的也较为容易收集的分类标准也可以直接纳入分类体系(Katuwal et al., 2016)。但是这类判别标准仍相对宽泛, 无法完全满足自闭症特异性对指标选取的精细化要求。未来研究可在已有研究成果的基础上, 一方面继续细化行为、量表、影像等指标适用范围, 按照分类性能(准确性、稳定性等)对具体指标进行规则排序; 另一方面应探究不同模态间如何有机结合、优势互补, 兼顾自闭症患儿的内部神经机制和外在行为表现, 更好地激发多模态数据隐藏的价值。

最后, 理论模型向现实应用的转化还未完全成熟。机器学习可以帮助家长或其他从业人员提前知晓疾病风险, 及早制定康复方案。但是现有研究仍停留在模型构建的理论阶段, 模型应用到实践中还需要慎重思考以下几个问题。一是研究样本对总体的代表性。已有研究模型中选用的被试通常是经过选择的, 多数为高风险自闭症家族人群, 自闭症所占比重远超典型发展人群正常值(20%~33%; 1%), 评估结果若直接应用到普通人群中是否能够兼容(Rahman et al., 2020; Thabtah, 2019); 二是模型的适用性和可推广性。受限于研究可行性因素, 目前超过 1/3 的研究样本量不足 300 人, 可能存在采样偏差, 且很少有研究报告成果在不同文化背景、地理区域被试群体中的验证, 理论模型的泛化能力还不能确定(Liaqat et al., 2021; Mazumdar et al., 2021)。三是算法的准确度和稳定性有待改善。当前模型仍多采用为传统机器学习算法, 部分模型分类准确率尚不足 60%, 稳定性也难以满足大范围施测要求。四是模型功能较为单一。当前研究成果只能对个体是否为自闭症做简单的二分类判断, 极少直接应用于诊断

(Millar et al., 2019)。对于计算模型应用于预测和评估过程管理的功能也是浅尝辄止、不够深入。五是应用受众对模型的认可度仍需提高。当前ASD Tests、Cognoa等操作简便的智能识别软件已经可以轻松地在手机端下载获得，上传视频或完成测试题目就可以快速获得筛查结果，但是绝大多数家长、教师甚至是相关专业从业人员并不了解这些软件。即使家长等相关人员知晓这些施测工具，关系到识别尤其是“诊断”结果，对于工具的认可和接受程度可能也难以达到研究者预期。对比传统方法，专业医师经过长时间观察和家长访谈，结合高度结构化量表慎重考量得出结果，仅 ADI-R 量表一项施测时间就长达 2.5 小时，过程蕴含大量自然真实环境中的人际互动线索(Abbas et al., 2020)。机器学习模型在缩短施测时间的同时能否得到真实有效的数据，预先设定的训练数据是否能完全覆盖测试对象，评估结果如何解释，评估结果如何应用于后续干预与康复，这些都是家长或其他应用受众会关注的问题。随着机器学习技术的不断成熟，上述问题更深入的思考与论证将有助于加快推进理论模型向实践转化的步伐。未来研究可在丰富研究对象和优化机器学习算法的基础上，充分调研不同受众的实际需求与困扰，有针对性的解决问题，同时增加对理论研究成果的宣导，在实践中不断加深理论模型的渗透与推广。

参考文献

*为纳入系统分析的文献

陈光华，陶冠澎，翟璐煜，白学军．(2022). 自闭症谱系障碍的早期筛查工具．*心理科学进展, 30*(4), 738−760.

陈顺森，白学军，张日昇．(2011). 自闭症谱系障碍的症状、诊断与干预．*心理科学进展, 19*(1), 60−72.

曹正凤．(2014). *随机森林算法优化研究* (博士学位论文). 首都经济贸易大学，北京.

崔建伟，赵哲，杜小勇．(2021). 支撑机器学习的数据管理技术综述．*软件学报, 32*(3), 604−621.

董健宇，韦文棋，吴珂，妮娜，王粲霏，付莹，彭歆．(2020). 机器学习在抑郁症领域的应用．*心理科学进展, 28*(2), 266−274.

顾曰国．(2013). 论言思情貌整一原则与鲜活话语研究——多模态语料库语言学方法．*当代修辞学*, (6), 1−19.

郭利明，杨现民，段小莲，邢蓓蓓．(2019). 人工智能与特殊教育的深度融合设计．*中国远程教育*, (8), 10−19+92−93.

韩保祯．(2018). *基于多模态信息融合的脑网络研究及其应用* (硕士学位论文). 电子科技大学，成都.

胡越，罗东阳，花奎，路海明，张学工．(2019). 关于深度学习的综述与讨论．*智能系统学报, 14*(1), 1−19.

黄碧玉．(2018). *社交机器人对孤独症儿童社交行为诱导效果研究* (硕士学位论文). 浙江工业大学，杭州.

靳宇倡，张政，郑佩璇，安俊秀．(2022). 远程心理健康服务：应用、优势及挑战.*心理科学进展, 30*(1), 141−156.

李航．(2012). *统计学习方法*. 北京：清华大学出版社.

刘笑晗，陈明隆，郭静．(2022). 机器学习在儿童创伤后应激障碍识别及转归预测中的应用．*心理科学进展, 30*(4), 1−12.

徐建平，张厚粲．(2005). 质性研究中编码者信度的多种方法考察．*心理科学, 28*(6), 1430−1432.

杨剑锋，乔佩蕊，李永梅，王宁．(2019). 机器学习分类问题及算法研究综述．*统计与决策, 35*(6), 36−40.

余明华，冯翔，祝智庭．(2017). 人工智能视域下机器学习的教育应用与创新探索．*远程教育杂志, 35*(3), 11−21.

*Abbas, H., Garberson, F., Glover, E., & Wall, D. P. (2018). Machine learning approach for early detection of autism by combining questionnaire and home video screening. *Journal of the American Medical Informatics Association, 25*(8), 1000−1007.

*Abbas, H., Garberson, F., Liu-Mayo, S. Glover, E., & Wall, D. P. (2020). Multi-modular AI approach to streamline autism diagnosis in young children. *Scientific Reports, 10*(1), 1−8.

*Abdolzadegan, D., Moattar, M. H., & Ghoshuni, M. (2020). A robust method for early diagnosis of autism spectrum disorder from EEG signals based on feature selection and DBSCAN method. *Biocybernetics and Biomedical Engineering, 40*(1), 482−493.

*Abraham, A., Milham, M. P., Martino, A. D., Craddock, R. C., Samaras, D., Thirion, B., & Varoquaux, G. (2017). Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example. *NeuroImage, 147*(15), 736−745.

*Achenie, L. E. K., Scarpa, A., Factor, R. S., Wang, T., Robins, D. L., & McCrickard, D. S. (2019). A machine learning strategy for autism screening in toddlers. *Journal of Developmental and Behavioral Pediatrics, 40*(5), 369−376.

Autism Brain Imaging Data Exchange. (2017). Retrieved July 3, 2021, from http://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html

*Bahado-Singh, R. O., Vishweswaraiah, S., Aydas, B., Mishra, N. K., Yilmaz, A., Guda, C., & Radhakrishna, U. (2019). Artificial intelligence analysis of newborn leucocyte epigenomic markers for the prediction of autism. *Brain Research, 1724*, 146457.

*Bone, D., Bishop, S., Black, M. P., Goodwin, M. S., Lord, C., & Narayanan, S. S. (2016). Use of machine learning to

improve autism screening and diagnostic instruments: Effectiveness, efficiency, and multi-instrument fusion. *Journal of Child Psychology and Psychiatry, 57*(8), 927−937.

Bone, D., Goodwin, M. S., Black, M. P., Lee, C., Audhkhasi, K., & Narayanan, S. (2015). Applying machine learning to facilitate autism diagnostics: Pitfalls and promises. *Journal of Autism and Developmental Disorders, 45*(5), 1121−1136.

*Bosl, W. J., Tager-Flusberg, H., & Nelson, C. A. (2018). EEG analytics for early detection of autism spectrum disorder: A data-driven approach. *Scientific Reports, 8*(1), 6828.

*Ben-Sasson, A., & Yom-Tov, E. (2016). Online concerns of parents suspecting autism spectrum disorder in their child: Content analysis of signs and automated prediction of risk. *Journal of Medical Internet Research, 18*(11), e300.

*Bussu, G., Jones, E. J., Charman, T., Johnson, M. H., & Buitelaar, J. K. (2018). Prediction of autism at 3 years from behavioural and developmental measures in high-risk infants: A longitudinal cross-domain classifier analysis. *Journal of Autism and Developmental Disorders*, *48*(7), 2418−2433.

*Caly, H., Rabiei, H., Coste-Mazeau, P., Hantz, S., Alain, S., Eyraud, J. L., … Ben-Ari, Y. (2021). Machine learning analysis of pregnancy data enables early identification of a subpopulation of newborns with ASD. *Scientific Reports*, *11*(1), 6877.

Choi, G. B., Yim, Y. S., Wong, H., Kim, S., Kim, H., Kim, S. V., … Huh, J. R. (2016). The maternal interleukin-17a pathway in mice promotes autism-like phenotypes in offspring. *Science*, *351*(6276), 933−939.

Cloarec, R., Riffault, B., Dufour, A., Rabiei, H., Gouty-Colomer, L. A., Dumon, C., … Ben-Ari, Y. (2019). Pyramidal neuron growth and increased hippocampal volume during labor and birth in autism. *Science Advances*, *5*(1), eaav0394.

Diehl, J. J., Schmitt, L. M., Villano, M., & Crowell, C. R. (2012). The clinical use of robots for individuals with autism spectrum disorders: A critical review. *Research in Autism Spectrum Disorders, 6*(1), 249−262.

*Dickinson, A., Daniel, M., Marin, A., Gaonkar, B., Dapretto, M., McDonald, N. M., & Jeste, S. (2020). Multivariate neural connectivity patterns in early infancy predict later autism symptoms. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *6*(1), 59−69.

*Dong, H. Y., Chen, D., Zhang, L., Ke, H. J., & Li, X. L. (2021). Subject sensitive EEG discrimination with fast reconstructable CNN driven by reinforcement learning: A case study of ASD evaluation. *Neurocomputing*, *449*, 136−145.

Dubreuil-Vall, L., Ruffini, G., & Camprodon, J. A. (2020). Deep learning convolutional neural networks discriminate adult ADHD from healthy individuals on the basis of event-related spectral EEG. *Frontiers in Neuroscience, 14*, 251.

*Duda, M., Daniels, J., & Wall, D. P. (2016). Clinical evaluation of a novel and mobile autism risk assessment. *Journal of Autism and Developmental Disorders*, *46*(6), 1953−1961.

*Duda, M., Haber, N., Daniels, J., & Wall, D. P. (2017). Crowdsourced validation of a machine-learning classification system for autism and ADHD. *Translational Psychiatry*, *7*(5), e1133.

*Duda, M., Ma, R., Haber, N., & Wall, D. P. (2016). Use of machine learning for behavioral distinction of autism and ADHD. *Translational Psychiatry*, *6*(2), e732.

Elsabbagh, M., Mercure, E., Hudry, K., Chandler, S., Pasco, G., Charman, T., … the BASIS Team. (2012). Infant neural sensitivity to dynamic eye gaze is associated with later emerging autism. *Current Biology*, *22*(4), 338−342.

*Emerson, R. W., Adams, C., Nishino, T., Hazlett, H, C., Wolff, J. J., Zwaigenbaum, L., … Piven, J. (2017). Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age. *Science Translational Medicine*, *9*(393). https://doi.org/10.1126/scitranslmed.aag2882

*Eni, M., Dinstein, I., Ilan, M., Menashe, I., Meiri, G., & Zigel, Y. (2020). Estimating autism severity in young children from speech signals using a deep neural network. *IEEE Access*, *8*, 139489−139500.

*Epalle, T. M., Song, Y. Q., Liu, Z., & Lu, H. (2021). Multi-atlas classification of autism spectrum disorder with hinge loss trained deep architectures: ABIDE I results. *Applied Soft Computing*, *107*(3), 107375.

*Eslami, T., Mirjalili, V., Fong, A., Laird, A. R., & Saeed, F. (2019). ASD-DiagNet: A hybrid learning approach for detection of autism spectrum disorder using fMRI data. *Frontiers in Neuroinformatics, 13*. https://doi.org/10.3389/fninf.2019.00070

*Eugenia, C., Alessandra, R., Letizia, P., Spera, G., Bosco, P., Biagi, L., … Calderoni, S. (2020). Autism spectrum disorder and childhood apraxia of speech: Early language-related hallmarks across structural MRI study. *Journal of Personalized Medicine, 10*(4), 275.

Falkmer, T., Anderson, K., Falkmer, M., & Horlin, C. (2013). Diagnostic procedures in autism spectrum disorders: A systematic literature review. *European Child and Adolescent Psychiatry*, *22*(6), 329−340.

Food and Drug Administration. (2021). *FDA authorizes marketing of diagnostic aid for autism spectrum disorder*. Retrieved October 21, 2021, from https://www.fda.gov/

news-events/press-announcements/fda-authorizes-marketing-diagnostic-aid-autism-spectrum-disorder

*Feczko, E., Balba, N., Miranda-Dominguez, O., Cordova, M., Karalunas, S. L., Irwin, L., ... Fair, D. A. (2017). Subtyping cognitive profiles in autism spectrum disorder using a random forest algorithm. *Neuroimage, 172*, 674−688.

Geng, X., Kang, X., & Wong, P. C. M. (2020). Autism spectrum disorder risk prediction: A systematic review of behavioral and neural investigations. *Progress in Molecular Biology and Translational Science, 173*, 91−137.

*Goel, N., Grover, B., Anuj, Gupta, D., Khanna, A., & Sharma, M. (2020). Modified grasshopper optimization algorithm for detection of autism spectrum disorder. *Physical Communication, 41*, 101115.

*Gök, M. (2018). A novel machine learning model to predict autism spectrum disorders risk gene. *Neural Computing and Applications, 31*(4), 6711−6717.

*Grossi, E., Olivieria, C., & Buscema, M. (2017). Diagnosis of autism through EEG processed by advanced computational algorithms: A pilot study. *Computer Methods and Programs in Biomedicine*, *142*, 73−79.

Guinchat, V., Chamak, B., Bonniau, B., Bodeau, N., Perisse, D., Cohen, D., & Danion, A. (2012). Very early signs of autism reported by parents include many concerns not specific to autism criteria. *Research in Autism Spectrum Disorders, 6*(2), 589−601.

*Guo, X. Y., Dominick, K. C., Minai, A. A., Li, H. L., Erickson, C. A., & Lu, L. J. (2017). Diagnosing autism spectrum disorder from brain resting-state functional connectivity patterns using a deep neural network with a novel feature selection method. *Frontiers in Neuroscience*, *11*, 460.

Hazlett, H. C., Gu, H., Munsell, B. C., Kim, S. H., Styner, M., Wolff, J. J., … The IBIS Network. (2017). Early brain development in infants at high risk for autism spectrum disorder. *Nature*, *542*(7641), 348−351.

Hosseinzadeh, M., Koohpayehzadeh, J., Bali, A. O., Rad, F. A., Souri, A., Mazaherinezhad, A., ... Bohlouli, M. (2021). A review on diagnostic autism spectrum disorder approaches based on the internet of things and machine learning. *The Journal of Supercomputing, 77*(3), 2590−2608.

Hyde, K. K., Novack, M. N., LaHaye, N., Parlett-Pelleriti, C., Anden, R., Dixon, D. R., & Linstead, E. (2019). Applications of supervised machine learning in autism spectrum disorder research: A review. *Review Journal of Autism and Developmental Disorders, 6*(2), 128−146.

*Itani, S., & Thanou, D. (2021). Combining anatomical and functional networks for neuropathology identification: A case study on autism spectrum disorder. *Medical Image Analysis, 69*, 101986.

*Jun, E., Kang, E., Choi, J., & Suk, H. (2019). Modeling regional dynamics in low-frequency fluctuation and its application to autism spectrum disorder diagnosis. *NeuroImage, 184*, 669−686.

*Kang, J. N., Han, X. Y., Song, J. J., Niu, Z. K., & Li, X. L. (2020). The identification of children with autism spectrum disorder by SVM approach on EEG and eye-tracking data. *Computers in Biology and Medicine, 120*(24), 103722.

*Kanne, S. M., Carpenter, L. A., & Warren, Z. (2018). Screening in toddlers and preschoolers at risk for autism spectrum disorder: Evaluating a novel mobile-health screening tool. *Autism Research, 11*(7), 1038−1049.

*Katuwal, G. J., Baum, S. A., Cahill, N. D., & Michael, A. M. (2016). Divide and conquer: Sub-grouping of ASD improves ASD detection based on brain morphometry. *Plos One*, *11*(4), 1−24.

Klin, A., Lang, J., Cicchetti, D. V., & Volkmar, F. R. (2000). Brief report: Interrater reliability of clinical diagnosis and DSM-IV criteria for autistic disorder: Results of the DSM-IV autism field trial. *Journal of Autism and Developmental Disorders*, *30*(2), 163−167.

Koumpouros, Y., & Kafazis, T. (2019). Wearables and mobile technologies in autism spectrum disorder interventions: A systematic literature review. *Research in Autism Spectrum Disorders, 66*, 101405.

*Lai, M., Lee, J., Chiu, S., Charm, J., So, W. Y., Yuen, F. P., … Zee, B. (2020). A machine learning approach for retinal images analysis as an objective screening method for children with autism spectrum disorder. *Eclinical Medicine*, *28*, 100588.

*Levy, S., Duda, M., Haber, N., Wall, D. P. (2017). Sparsifying machine learning models identify stable subsets of predictive features for behavioral detection of autism. *Molecular Autism, 8*(1), 65.

*Li, G. Y., Lee, O., & Rabitz, H. (2018). High efficiency classification of children with autism spectrum disorder. *Plos One, 13*(2), e0192867.

Li, X. W., La, R., Wang, Y., Niu, J. H., Zeng, S., Sun, S., & Zhu, J. (2019). EEG-based mild depression recognition using convolutional neural network. *Medical and Biological Engineering and Computing, 57*(6), 1341−1352.

*Li, Y. M., Mache, M. A., & Todd, T. A. (2020). Automated identification of postural control for children with autism spectrum disorder using a machine learning approach. *Journal of Biomechanics*, *113*, 110073.

*Liaqat, S., Wu, C., Duggirala, P. R., Cheung, S. S., Chuah, C. N., Ozonoff, S., & Young, G. (2021). Predicting ASD diagnosis in children with synthetic and image-based eye

gaze data. *Signal Processing Image Communication, 94*(3), 116198.

*Liu, W., Li, M., & Li, Y. (2016). Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework. *Autism Research, 9*(8), 888−898.

Maenner, M. J., Shaw, K. A., Bakian, A.V., Bilder, D. A., Durkin, M. S., Esler, A., … Cogswell, M. E. (2021). Prevalence and characteristics of autism spectrum disorder among children aged 8 years - Autism and developmental disabilities monitoring network, 11 sites, United States, 2018. *MMWR Surveillance Summaries*. *70*(11), 1−16.

*Mazumdar, P., Arru, G., & Battisti, F. (2021). Early detection of children with autism spectrum disorder based on visual exploration of images. *Signal Processing Image Communication, 94*(3), 116184.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., PRISMA Group, & for the PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ, 339*, b2535.

*Meera, S. S., Donovan, K., Wolff, J. J., Zwaigenbaum, L., Elison, J. T., Kinh, T., ... IBIS Network. (2020). Towards a data-driven approach to screen for autism risk at 12 months of age. *Journal of the American Academy of Child and Adolescent Psychiatry, 60*(8), 968−977

Millar, L., Mcconnachie, A., Minnis, H., Wilson, P., Thompson, L., Anzulewicz, A., … Delafield-Butt, J. (2019). Phase 3 diagnostic evaluation of a smart tablet serious game to identify autism in 760 children 3-5 years old in sweden and the united kingdom. *BMJ Open*, *9*(7), e026226.

Namdeo, A., & Singh, D. (2021). Challenges in evolutionary algorithm to find optimal parameters of SVM: A review. *Materials Today: Proceedings, 3*, 288.

*Negin, F., Ozyer, B., Agahian, S., Kacdioglu, S., & Ozyer, G. T. (2021). Vision-assisted recognition of stereotype behaviors for early diagnosis of autism spectrum disorders. *Neurocomputing*, *446*, 145−155.

Nogay, H. S., & Adeli, H. (2020). Machine learning (ML) for the diagnosis of autism spectrum disorder (ASD) using brain imaging. *Reviews in the Neurosciences, 31*(8), 825−841.

Oh, D. H., Kim, I. B., Kim, S. H., & Ahn, D. H. (2017). Predicting autism spectrum disorder using blood-based gene expression signatures and machine learning. *Clinical Psychopharmacology and Neuroscience: The Official Scientific Journal of the Korean College of Neuropsychopharmacology, 15*(1), 47−52.

Pagnozzi, A. M., Conti, E., Calderoni, S., Fripp, J., & Rose, S. E. (2018). A systematic review of structural MRI biomarkers in autism spectrum disorder: A machine learning perspective. *International Journal of Developmental Neuroscience, 71*(1), 68−82.

*Parikh, M. N., Li, H., & He, L. (2019). Enhancing diagnosis of autism with optimized machine learning models and personal characteristic data. *Frontiers in Computational Neuroscience, 13*, 9.

Perera, H., Jeewandara, K. C., Seneviratne, S., & Guruge, C. (2017). Culturally adapted pictorial screening tool for autism spectrum disorder: A new approach. *World Journal of Clinical Pediatrics*, *6*(1), 45−51.

Plitt, M., Barnes, K. A., & Martin, A. (2015). Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards. *NeuroImage Clinical, 7*(C), 359−366.

*Puerto, E., Aguilar, J., López, C., & Chávez, D. (2019). Using multilayer fuzzy cognitive maps to diagnose autism spectrum disorder. *Applied Soft Computing, 75*, 58−71.

*Qiu, N., Tang, C., Zhai, M., Huang, W., Weng, J., Li, C., … Ke, X. (2020). Application of the still-face paradigm in early screening for high-risk autism spectrum disorder in infants and toddlers. *Frontiers in Pediatrics*, *8*, 290.

Quaak, M., van de Mortel, L., Thomas, R. M., & van Wingen, G. (2021). Deep learning applications for the classification of psychiatric disorders using neuroimaging data: Systematic review and meta-analysis. *NeuroImage Clinical, 30*, 102584.

*Rahman, R., Kodesh, A., Levine, S. Z., Sandin, S., Reichenberg, A., & Schlessinger, A. (2020). Identification of newborns at risk for autism using electronic medical records and machine learning. *European Psychiatry. 63*(1), e22, 1−7.

Rahman, M. M., Usman, O. L., Muniyandi, R. C., Sahran, S., Mohamed, S., & Razak, R. A. (2020). A review of machine learning methods of feature selection and classification for autism spectrum disorder. *Brain Sciences, 10*(12), 1−23.

*Raj, S., & Masood, S. (2020). Analysis and detection of autism spectrum disorder using machine learning techniques. *Procedia Computer Science, 167*, 994−1004.

*Rakić, M., Cabezas, M., Kushibar, K., Oliver, A., & Lladó, X. (2020). Improving the detection of autism spectrum disorder by combining structural and functional MRI information. *NeuroImage: Clinical*, *25*, 102181.

Ramírez-Duque, A. A., Frizera-Neto, A., & Bastos, T. F. (2019). Robot-assisted autism spectrum disorder diagnostic based on artificial reasoning. *Journal of Intelligent and Robotic Systems*, *96*(2), 267−281.

*Retico, A., Gori, I., Giuliano, A., Muratori, F., & Calderoni, S. (2016). One-class support vector machines identify the

language and default mode regions as common patterns of structural alterations in young children with autism spectrum disorders. *Frontiers in Neuroscience, 10*(171), 306.

*Salari, V., Sherkatghanad, Z., Akhondzadeh, M., Salari, S., Zomorodi-moghadam, M., Abdar, M., & Khosrowabadi, R. (2019). Automated detection of autism spectrum disorder using convolutional neural network. *Frontiers in Neuroscience. 13*, 1325.

Sanders, S. J., Ercan-Sencicek, A. G., Hus, V., Luo, R., Murtha, M. T., Moreno-De-Luca, D., … State, M. W. (2011). Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron*, *70*(5), 863−885.

Sekaran, K., & Sudha, M. (2020). Predicting autism spectrum disorder from associative genetic markers of phenotypic groups using machine learning. *Journal of Ambient Intelligence and Humanized Computing, 12*(3), 3257−3270.

*Sen, B., Borle, N. C., Greiner, R., & Brown, M. R. G. (2018). A general prediction model for the detection of ADHD and autism using structural and functional MRI. *Plos One*, *13*(4), e0194856.

*Sewani, H., & Kashef, R. (2020). An autoencoder-based deep learning classifier for efficient diagnosis of autism. *Children, 7*(10), 0182.

*Shahamiri, S. R., & Thabtah, F. (2020). Autism AI: A new autism screening system based on artificial intelligence. *Cognitive Computation*, *12*, 766−777.

Sheldrick, R. C., Maye, M. P., & Carter, A. S. (2017). Age at first identification of autism spectrum disorder: An analysis of two US surveys. *Journal of the American Academy of Child and Adolescent Psychiatry*, *56* (4), 313−320.

Shic, F., Macari, S., & Chawarska, K. (2014). Speech disturbs face scanning in 6-month-old infants who develop autism spectrum disorder. *Biological Psychiatry*, *75*(3), 231−237.

*Tariq, Q., Daniels, J., Schwartz, J. N., Washington, P., Kalantarian, H., & Wall, D. P. (2018). Mobile detection of autism through machine learning on home video: A development and prospective validation study. *Plos Medicine, 15*(11), e1002705.

*Thabtah, F., Kamalov, F., & Rajab, K. (2018). A new computational intelligence approach to detect autistic features for autism screening. *International Journal of Medical Informatics, 117*(SEP.), 112−124.

*Thabtah, F. (2019). An accessible and efficient autism screening method for behavioural data and predictive analyses. *Health Informatics Journal, 25*(2), 1739-1755.

*Thabtah, F., Abdelhamid, N., & Peebles, D. (2019). A machine learning autism classification based on logistic regression analysis. *Health Information Science and Systems, 7*(1), 12.

*Thabtah, F., & Peebles, D. (2019). A new machine learning model based on induction of rules for autism detection. *Health Informatics Journal, 26*(1), 264−286.

Tyzio, R., Nardou, R., Ferrari, D. C., Tsintsadze, T., Shahrokhi, A., Eftekhari, S., … Ben-Ari, Y. (2014). Oxytocin-mediated gaba inhibition during delivery attenuates autism pathogenesis in rodent offspring. *Science*, *343*(6171), 675−679.

*Wan, G., Kong, X., Sun, B., Yu, S., Tu, Y., Park, J., … Kong, J. (2019). Applying eye tracking to identify autism spectrum disorder in children. *Journal of Autism and Developmental Disorders, 49*(1), 209−215.

Wolff, J. J., Gerig, G., Lewis, J. D., Soda, T., Styner, M. A., Vachet, C., … the IBIS Network. (2015). Altered corpus callosum morphology associated with autism over the first 2 years of life. *Brain*, *138*(7), 2046−2058.

*Wedyan, M., Al-Jumaily, A., & Crippa, A. (2019). Using machine learning to perform early diagnosis of autism spectrum disorder based on simple upper limb movements. *International Journal of Hybrid Intelligent Systems, 15*(4), 195−206.

Wolff, J. J., Gu, H., Gerig, G., Elison, J. T., Styner, M., Gouttard S., … the IBIS Network. (2012). Differences in white matter fiber tract development present from 6 to 24 months in infants with autism. *American Journal of Psychiatry*, *169*(6), 589−600.

*Wingfield, B., Miller, S., Yogarajah, P., Kerr, D., Gardiner, B., Seneviratne, S., … Coleman, S. (2020). A predictive model for paediatric autism screening. *Health Informatics Journal, 26*(4), 2538−2553.

*Yang, Y. (2020). A preliminary evaluation of still face images by deep learning: A potential screening test for childhood developmental disabilities. *Medical Hypotheses, 144*, 109978.

Application of machine learning in early identification and diagnosis of autistic children

HOU Tingting[1], CHEN Xiao[2], KONG Depeng[2], SHAO Xiujun[3], LIN Fengxun[1], LI Kaiyun[1]

([1] *School of Education and Psychology, University of Jinan, Jinan* 250022, *China*)

([2] *School of Educational Technology and Science, Zhejiang University of Technology, Hangzhou* 310023, *China*)

([3] *Qingdao Chenxing Experimental School, Qingdao* 266000, *China*)

Abstract: Early detection, early diagnosis and early intervention have been determined as the consensus of autistic children education and rehabilitation. However, both the limitations of traditional identification and diagnosis methods and the lack of professionals mostly caused the miss of the well-timed intervention in autistic children. Considering the advantages of machine learning (i.e., objectivity, accuracy, simplicity and flexibility), which has been gradually applied in different aspects of autistic education and rehabilitation, including early prediction, screening, diagnosis and evaluation, to improve this situation. Despite that, a series of limitations, such as research objects selection, classified data collection, theoretical model application, still exist. How to promote the construction of the tracking database of maternal and newborn pathophysiological information and establish a standardized model classification index system should be the focus of the future research. At the same time, researchers should pay more attention to the algorithm optimization and accelerate the transformation of the theoretical achievements of intelligent autism identification and diagnosis to practice.

Key words: machine learning, autism, early identification and diagnosis, systematic review

附录：

附录 1　机器学习算法与分类模型性能汇总表

序号	国家	作者	被试			主要数据指标	算法/方法	评价指标 Acc-/Sen-/ Spe-/AUC
			样本量	年龄	性别(男%)			
1	美国	Abbas et al.(2018)	162 (ASD+non)	18~72 个月	/	家长问卷，家庭视频	RF	–/100%/97.7%/99.7%
2	美国	Abbas et al.(2020)	375 (ASD+non)	18~72 个月	/	家长问卷，家庭视频，医生问卷	GBDT	–/90%/83%/ 92%
3	伊朗	Abdolzadegan et al.(2020)	34ASD, 11HC	3~12 岁	88.2; 54.5	脑电图	SVM, KNN	94.68%/100%/–/–
4	法国	Abraham et al.(2017)	871 (ASD+TC)	6~64 岁	83.5	fMRI	SVC, RR	66.9%/53.2%/78.3%/–
5	美国	Achenie et al.(2019)	14995 (ASD+non)	16~30 个月	46.51	个人信息, M-CHAT-R	fNN	99.72%/73.8%/99.9%/–
6	美国	Bahado-Singh et al.(2019)	14ASD, 10Controls	新生儿(24~79 小时)	/	白细胞 DNA	DL, RF, SVM, GLM, PAM, LDA	–/97.5%/100%/100%
7	以色列	Ben-Sasson & Yom-Tov(2016)	189 (LR+MR+HR)	1.25 个月~18 岁	84.7	父母问题	DT，最小二乘法	AUC = 82%
8	美国	Bone et al.(2016)	1264ASD, 462 non	4~55.1 岁	/	ADI-R, SRS	SVM	–/89.2%/59%/–
9	英国	Bosl et al.(2018)	99HR, 89LR	3 个月~3 岁	/	脑电图	KNN, RF, SVM	–/100%/100%/–
10	荷兰	Bussu et al.(2018)	232(HR+LR)	8~36 个月	50	ADOS, AOSI, VABS, MSEL	LSSVM	57.5%/79.6%/52.2%/71.3%
11	法国	Caly et al.(2021)	63ASD, 189NT	妊娠期	81; 48.7	超声波，生物测量	LASSO, DT, XGBoost	–/41%/96%/–
12	中国	Dong et al.(2021)	86ASD, 89TD	3~6 岁	/	脑电图	CNN	Acc– = 100%
13	美国	Duda & Ma et al.(2016)	2775ASD, 150ADHD	13~163 个月	83.9; 62	SRS	DT, RF, SVC, LR, LASSO, LDA	AUC = 96.5%
14	美国	Duda, Daniels, & Wall (2016)	222 (ASD+non)	16 个月~17 岁	76.1	ADI-R	ADTree	–/89.9%/79.7%/–
15	美国	Duda et al.(2017)	200ASD, 114ADHD	2~17 岁	80.6; 65.5	SRS	ENet, LR, SVM, LDA, RR	AUC = 89%
16	美国	Dickinson et al.(2020)	36HR, 29LR	3~18 个月	/	脑电图	SVM	$r = 0.76$
17	美国	Emerson et al.(2017)	59HR	6~24 个月	69.5	fcMRI	SVM	96.6%/81.8%/100%/–
18	以色列	Eni et al.(2020)	56ASD, 6TD, 10Other	av4 岁;av3 岁;av4 岁	87.5;83.3;90	语音特征	MLR, SVR, FC-DNN, CNN	$r = 0.72$(ADOS)
19	中国	Epalle et al.(2021)	506ASD, 532TC	av10~35 岁	87.7; 82.1	fMRI	MISO-DNN, PNN	Acc- = 79.13%
20	美国	Eslami et al.(2019)	550ASD, 485HC	av10~34 岁	81.5; 88.7	fMRI	SMOTE, SVM, RF	68.2%/66.7%/69.4%/74.4%
21	意大利	Eugenia et al.(2020)	26ASD, 24CAS, 18TD	34~74 个月	77; 75; 72	sMRI	SVM	AUC = 76% (ASD vs TD); AUC = 64% (ASD vs CAS)
22	美国	Feczko et al.(2017)	47ASD, 58TD	9~13 岁	80.85; 46.55	FcMRI	FRF	72.7%/80.7%/63.1%/–

续表

序号	国家	作者	被试			主要数据指标	算法/方法	评价指标 Acc-/Sen-/ Spe-/AUC
			样本量	年龄	性别(男%)			
23	土耳其	Gök (2018)	336ASD	8 周~40 岁	/	lncRNA 基因	NB，贝叶斯网络, KNN, LR, RF, SVM	83.6%/90.2%/98.8%/83.9%
24	印度	Goel et al.(2020)	1100ASD	"C&T&A"	/	行为特征，个人特征	MGOA, GOA, RF, KNN, NB, LR	100%/100%/100%/–
25	意大利	Grossi et al.(2017)	15ASD, 10TD	7~14 岁	86.7; 40	脑电图	TwIST, MSROM, MSE, PCA, ANN	Acc- = 100%
26	美国	Guo et al.(2017)	55ASD, 55TD	av12.7 岁, av14.2 岁	83.64; 69.09	fMRI	DNN	Acc- = 86.36%
27	比利时	Itani & Thanou(2021)	C: 201ASD, 251NT; A: 63ASD, 67NT	6~18 岁	/	fMRI	DT	Acc– = 74.8%
28	韩国	Jun et al.(2019)	121ASD, 171TD	6.5~39.1 岁	84.3; 74.27	fMRI	SVM	84.64%/80%/87.5%/84.5%
29	中国	Kang et al.(2020)	49ASD, 48TD	3~6 岁	79.6; 75	脑电图，眼动	SVM	85.44%/–/–/93%
30	美国	Kanne et al.(2018)	164ASD, 66non	18~72 个月	81.1; 75.8	家长问卷，家庭视频	ML	75%/75%/74%72.9%
31	美国	Katuwal et al.(2016)	361ASD, 373TDC	av18 岁	100	sMRI	RF	≥60/–/–/92%
32	中国	Lai et al.(2020)	46ASD, 24NC	7~20 岁	73.9;70.8	视网膜图像	SVM, ARIA	–/95.7%/91.3%/97.4%
33	美国	Levy et al.(2017)	1319ASD, 70non (M2); 2870ASD, 273non (M3)	av5 岁, av7 岁(M2); av9 岁(M3)	80.4; 58.6 (M2); 84.3; 59.2 (M3)	ADOS	LR, LASSO, RR, Enet, L1LR, L2LR, LDA	AUC = 93, 3% (M2); 95% (M3)
34	美国	Li et al.(2018)	83ASD, 76NEU	"C"	/	代谢物	SVM	Acc– > 96%
35	美国	Liaqat et al.(2021)	14ASD, 14TD	5~12 岁	/	合成扫视图像，图像	MLP, CNN, LSTM	67.23%/68%/66%/73%
36	美国	Li et al.(2020)	25ASD, 25TD	5~12 岁	44; 60	姿势控制数据	NB, LDA, KNN, DT, SVM, RF	90%/82.6%/100%, /–
37	中国	Liu et al.(2016)	29ASD, 58TD	4~11 岁	86.2; 86.2	眼动数据	SVM, KNN	88.51%/93.1%/86.21%/89.63%
38	意大利	Mazumdar et al.(2021)	14ASD, 14TD	"C"	/	眼动数据	DT, NB, SVM, KNN, EL	68.5%/68%/50%/-
39	美国	Meera et al.(2020)	237HR, 72LR	12 个月	/	父母问卷	RF	–/71%/72%/71%
40	法国	Negin et al.(2021)	108ASD	"C"	/	视觉单词，刻板行为视频	MLP, SVM, GNB, CNN, LSTM	79.28%/–/–/77.68%
41	美国	Parikh et al.(2019)	421ASD, 430non	av17 岁	88; 82	个人特征	KNN, SVM, DT, LR, NN, RF	AUC = 64.6%
42	美国	Puerto et al.(2019)	150ASD, 150NC	2~12 岁	74.7	ADI-R, ADOS	MFCM, SVM, RF, NB	Acc- = 86.5%
43	中国	Qiu et al.(2020)	45HR-ASD, 43TD	8~23 个月	88.9; 74.4	社会行为	SVM, NB, RF	Acc– = 81.18(FF); 83.35(SF)
44	英国	Rahman et al.(2020)	1397 ASD, 94741 non	0~18 岁	/	父母电子病历	LR, ANN/MLP, RF	96.5%/35.4%/99.3%/72.7%

续表

序号	国家	作者	被试			主要数据指标	算法/方法	评价指标 Acc-/Sen-/ Spe-/AUC
			样本量	年龄	性别(男%)			
45	西班牙	Rakić et al.(2020)	368ASD, 449Controls	7~64 岁	/	sMRI, fMRI	栈式自动编码器, MLP	85.06%/82%/89%/–
46	印度	Raj & Masood (2020)	292C, 104T, 704A	“C&T&A”	/	个人特征	NB, SVM, LR, ANN, KNN, NN, CNN	98.3%/96.79%/100%/–
47	意大利	Rrtico et al.(2016)	41ASD, 40Controls	22~72 个月	51.2; 50	MRI	SVM	AUC = 81.0(男)
48	新西兰	Shahamiri & Thabtah (2020)	6075(ASD+non)	“I&C&T&A”	58	AQ-10, Q-CHAT-10	CNN, C4.5, BN, Ridor, DL	97.95%/98.63%/96.53%/–
49	伊朗	Salari et al.(2019)	505ASD, 530TC	av10~33.7 岁	87.7; 82.1	fMRI	CNN, SVM, KNN, RF	70.22%/77.46%/61.82%/74.86%
50	加拿大	Sen et al.(2018)	1783	av0~30 岁	72.6	sMRI, fMRI	SVM, PCA	Acc- = 67.3%; 64.2%
51	加拿大	Sewani et al.(2020)	539ASD, 573TC	“C”	/	fMRI	SVM, RF, KNN, CNN	84.05%/-80%/84.1%/78%
52	新西兰	Thabtah et al.(2018)	707HR, 393LR	“C&T&A”	/	AQ-10	RIPPER, C4.5	92%/90.7%/93%/–
53	新西兰	Thabtah (2019)	707HR, 393LR	“C&T&A”	/	AQ-10	NB, LR	99.85%/99.9%/99.7%/–
54	新西兰	Thabtah et al.(2019)	248T, 1118A	“T&A”	57.5; 53.3	AQ-10	IG, 卡方检验	97.58%/97.6%/95.86%/–
55	新西兰	Thabtah & Peebles (2019)	292C, 104T, 704A	av29.2 岁	/	AQ-10	RML, RIPPER, Bagging, CART, C4.5, Prism	90%/91%/92%/–
56	美国	Tariq et al.(2018)	116ASD, 46TD	av5 岁, av3 岁	64.6	ADOS, ADI-R	ADTree, SVM, LR, 线性 SVM	96.2%/100%/77.3%/94%
57	中国	Wan et al.(2019)	37ASD, 37TD	4~6 岁	89.2; 73	眼动数据	SVM	85.1%/86.5%/83.8%/–
58	澳大利亚	Wedyan et al.(2019)	15ASD, 15TG	2~4 岁	80; 86.7	上肢运动	LDA, SVM	93.8%/91%/98.1%/–
59	英国	Wingfield et al.(2020)	195ASD, 33non	av2 岁	80; 84.8	PAAS 核对表	RF, Adaboost, DT, NN, NB	–/96%/97%/98%
60	中国	Yang (2020)	908ASD	“C”	/	自闭症面孔图片	CNN, SVM, DL	90%/96%/85%/93.8%

注：评价指标参数采用重点研究或最佳分类器的最优参数(比较顺序为 Acc-, AUC, Sen-, Spe-)；国家以第一作者所在国家为准。

Non：非 ASD; C (Children)：儿童; T (Teenager)：青少年; A (Adult)：成人; av (Average)：平均年龄; TD/TC/TG/HC/NC/NT/NEU (typically developing/ tytical control / typically Growing/ healthy control/normal control / neurotypical/ neurotypical controls)：典型发展儿童; MR (Medium-risk)：中风险; SVC/SVR (Support Vector Classification/ Regression)：支持向量分类/回归; LSSVM (Least Squares SVM)：最小二乘支持向量机; RML (Rules-Machine Learning)：规则机器学习; CART (Classification and Regression Tree)：分类回归树; GLM (Generalized Linear Model)：广义线性模型; TwIST (Two-Step Iterative Shrinkage/Thresholding)：两步迭代收缩; PCA (Principal Component Analysis))：主要成分分析; MSROM (multi scale ranked organizing maps)：多比例尺排列组织图; MSE (Multi scale entropy)：多尺度熵; FC-DNN (fully connected deep neural networks)：全连接的深度神经网络; SMOTE (Synthetic Minority Over-sampling Technique)：人工少数类过采样; NN (neural network)：神经网络; BLSTM NN (bidirectional long short-term memory neural network)：双向长短时记忆神经网络; IG (Information Gain)：信息增益; ARIA (Automatic retinal image analysis)：视网膜图像自动分析; FRF (Functional Random Forest)：功能随机森林; GNB (Gaussian Naive Bayes)：高斯朴素贝叶斯; MGOA (Modified Grasshopper Optimization Algorithm)：改进的蚱蜢优化算法; MFCM (Multilayer Fuzzy Cognitive Maps)：多层模糊认知地图; LSTM (Long Short-Term Memory)：长短期记忆网络; MISO-DNN (multi-input single-output deep neural network)：多输入单输出深度神经网络; PNN (probabilistic neural network)：概率神经网络; PAAS (pictorial autism assessment schedule)：自闭症图片评估程序.